Developmental Biology **220,** 307–321 (2000)
doi:10.1006/dbio.2000.9648, available online at http://www.idealibrary.com on IDEAL®

Dichaete Reveal New Functions in Embryonic Brain and Hindgut Development

Natalia Sánchez-Soriano[1] and Steven Russell[2]

Department of Genetics, University of Cambridge, Downing Street, Cambridge CB2 3EH, United Kingdom

***Sox* domain proteins encompass a conserved family of transcriptional regulators that are implicated in a variety of developmental processes in eukaryotes from worm to man. The *Dichaete* gene of *Drosophila* encodes a group B Sox protein related to mammalian Sox1, -2, and -3 and, like these proteins, it is widely and dynamically expressed throughout embryogenesis. In order to unravel new *Dichaete* functions, we characterized the organization of the *Dichaete* gene using a combination of regulatory mutant alleles and reporter gene constructs. *Dichaete* expression is tightly controlled during embryonic development by a complex of regulatory elements distributed over 25 kb downstream and 3 kb upstream of the transcription unit. A series of regulatory alleles which affect tissue-specific domains of *Dichaete* were used to demonstrate that *Dichaete* has functions in addition to those during segmentation and midline development previously described. First, *Dichaete* has functions in the developing brain. A specific group of neural cells in the tritocerebrum fails to develop correctly in the absence of *Dichaete,* as revealed by reduced expression of *labial, zfh-2, wingless,* and *engrailed.* Second, *Dichaete* is required for the correct differentiation of the hindgut. The *Dichaete* requirement in hindgut morphogenesis is, in part, via regulation of *dpp,* since ectopically supplied *dpp* can rescue *Dichaete* phenotypes in the hindgut. Taken together, there are now four distinct *in vivo* functions described for *Dichaete* that can be used as models for context-dependent comparative studies of *Sox* function.**

***Key Words: Drosophila; Dichaete; Sox;* brain; hindgut.**

INTRODUCTION

Sox genes encode a group of transcriptional regulators related to the mammalian testis determining factor, SRY, by virtue of their 79 amino acid HMG-box DNA-binding domain (Gubbay *et al.*, 1990; Sinclair *et al.*, 1990; Pevny and Lovell-Badge, 1997). Outside of the HMG-domain *Sox* genes are diverse; over 20 have been identified in a variety of eukaryotes ranging from worms to man and these are divided into six subclasses, A to F, based on homology both within and outside of the HMG-box (Pevny and Lovell-Badge, 1997). Sox proteins bind to DNA in the minor groove with moderate sequence specificity and are suggested to have an architectural role (i.e., DNA bending) or a role as traditional transcriptional regulators, either alone or in conjunction with other transcription factors (Pevny and Lovell-Badge, 1997). Some studies suggest that the same Sox protein can act in a variety of different ways. For example, Sox2 interacts with the POU transcription factor OCT3 in order to activate transcription of the FGF4 enhancer, however, it can also repress OCT3-mediated activation of the i-opn enhancer (Ambrosetti *et al.*, 1997; Botquin *et al.*, 1998). In addition, Sox2 can activate the DC5 enhancer of the α-crystallin gene together with the unidentified factor αEF3 (Kamachi *et al.*, 1995). Thus, Sox proteins can function in different ways depending on the respective cellular and molecular contexts.

Many insights into Sox protein functions come from *in vitro* studies, but to date little is known about their functions in the living organism. All of the *Sox* genes characterized to date show restricted temporal and tissue-specific expression profiles, suggesting that they may play specific regulatory roles in development. For some *Sox* genes the

[1] Present address: Institut fuer Genetik–Zellbiologie, Universitaet Mainz, Becherweg 32, D-55128 Mainz, Germany. Fax: +49/6131/394584. E-mail: sanchez@mail.uni-mainz.de.

[2] To whom correspondence should be addressed. Fax: +44 1223 333992. E-mail: sr120@mole.bio.cam.ac.uk.

0012-1606/00 $35.00

analysis of mutant phenotypes supports this view; for example, *SRY* mutant males fail to develop testis and *SOX9* mutant patients have defects in both bone morphogenesis and testicular development. These phenotypes correlate with a loss of *Sox* function in the anlage of the affected tissues (Foster *et al.*, 1994; Gubbay *et al.*, 1990). However, for other *Sox* genes, *in vivo* studies of their function have been less informative. Vertebrate *Sox1* is widely expressed in the developing central nervous system (CNS) of mouse and chicken. Although cell culture experiments indicate that this expression is required for neural determination, the development of the CNS in *Sox1* knockout mice is nearly normal (Kamachi *et al.*, 1998; Nishiguchi *et al.*, 1998; Pevny *et al.*, 1998). The lack of phenotype may be due to functional redundancy since two closely related genes (*Sox2* and *Sox3*) are widely coexpressed with *Sox1* in the developing CNS and all three are interchangeable to an extent when assayed in cell culture (Collignon *et al.*, 1996; Pevny *et al.*, 1998). Another problem with the *in vivo* analysis of *Sox* genes is that of pleiotropy, where early defects obscure the analysis of functions later in development. For example, *Sox2* is expressed early in mouse development in the inner cell mass and *Sox2* null mutant embryos die just after implantation. A potential function of later *Sox2* expression in the developing nervous system has therefore not been studied *in vivo* (Pevny *et al.*, 1998).

The *Drosophila Dichaete* gene is a group B *Sox* gene related to vertebrate *Sox1*, *2*, and *3* (Nambu and Nambu, 1996; Russell *et al.*, 1996). We have previously shown that *Dichaete* and *Sox2* are functionally conserved since mouse *Sox2* efficiently rescues *Dichaete* mutant phenotypes in flies (Sánchez-Soriano and Russell, 1998). Like mammalian *Sox* genes, *Dichaete* is dynamically expressed throughout embryogenesis in a variety of tissues and we have shown *in vivo* functions for *Dichaete* in some of the tissues in which it is expressed. First, during the early events of embryonic segmentation and, second, in the development of specific glia cells in the midline of the CNS (Nambu and Nambu, 1996; Russell *et al.*, 1996; Sánchez-Soriano and Russell, 1998). However, since further strong *Dichaete* expression domains exist (e.g., throughout the developing nervous system and in the gut) *Dichaete* might have further potential functions during embryonic development. We are interested in unraveling such additional developmental roles in order to have access to a set of *in vivo* models with different molecular contexts in which Sox protein functions can be studied comparatively.

Here we describe further *in vivo* functions of *Dichaete*. As a prerequisite to this study we have characterized a set of hypomorphic *Dichaete* alleles which uncover a 30-kb region containing regulatory elements responsible for the complex tissue-specific and temporal pattern of *Dichaete* expression. Making use of *Dichaete* alleles that eliminate *Dichaete* expression tissue specifically, we demonstrate that *Dichaete* is required for the differentiation of a restricted group of neural cells in the brain and for the correct development of the hindgut. Together with the previously described requirements in segmentation and midline development there are now four distinct *in vivo* roles for *Dichaete* which can be used for further investigations of *Sox* gene function.

MATERIALS AND METHODS

Molecular Biology

Dichaete cDNA and genomic clones are described in Russell *et al.*, (1996). Additional DNA from the *Dichaete* region was obtained from the Berkeley *Drosophila* Genome Project as P1 clone DS08449. Routine subcloning into plasmid vectors and Southern blot analysis was carried out using standard procedures (Sambrook *et al.*, 1989). The molecular breakpoints of aberrations were mapped with at least four different restriction enzymes and compared with progenitor chromosomes when available (all except D^1, D^3, and D^4). To identify regulatory elements, restriction fragments around the *Dichaete* transcription unit were cloned into the pP{CaSpeR-AUG-βgal} (Thummel *et al.*, 1988) vector and introduced into *y w* embryos using standard techniques (Karess, 1985).

In Situ Hybridization and Immunohistochemistry

Embryo staging was according to Campos-Ortega and Hartenstein (1997). *In situ* hybridization with digoxigenin-labeled cDNA probes was carried out with minor modifications to the procedure of Tautz and Pfeifle (1989). *hh* and *dpp* cDNA probes were a gift from J. de Celis. Antibody stainings were carried out essentially as described (Patel, 1994). Primary antibodies were detected with biotin-conjugated secondary antibodies and the ABC Elite kit (Vectastain) or with fluorescent secondary antibodies (Jackson ImmunoResearch). The following primary antibodies were used at the indicated dilutions: rabbit anti-*Dichaete* 1/2000 (Sánchez-Soriano and Russell, 1998), rabbit anti-β-galactosidase 1/10,000 (Cappel), mouse anti-Wingless 1/50 (a gift from S. Cohen; Diaz-Benjumea and Cohen, 1994), mouse anti-αSpectrin (Developmental Studies Hybridoma Bank, University of Iowa, Iowa City, IA), mouse anti-Engrailed 1/10 (Developmental Studies Hybridoma Bank), mouse anti-Zfh-2 1/300 (a gift from Z. C. Lai; Lai *et al.*, 1991), and rat anti-Labial 1/100 (a gift from M. Bienz; Diederich *et al.*, 1989).

Drosophila Stocks

Drosophila stocks were maintained on standard yeasted cornmeal–agar food at 25°C. Mutant nomenclature follows FlyBase conventions (FlyBase, 1998). A brief description of the *Dichaete* alleles used in this work is given in Table 1. A full description of the mutant alleles and other aberrations in the *Dichaete* region can be found in FlyBase. Throughout D^3 was used as a representative null allele and *Df*(3*L*) *fz*-Gs1a as a deficiency. *UAShh* and *UASdpp* were gifts from J. deCelis and *enGAL4* was a gift from A. Hidalgo.

RESULTS

Dichaete Mutant Alleles with Chromosomal Rearrangements Outside of the Transcription Unit

Dichaete null alleles have early segmentation defects which cause secondary defects in many tissues during later

TABLE 1

Relevant Genetics of the Alleles Used in This Work, Full Descriptions, and Relevant References May Be Found in FlyBase

Allele	Cytology	Molecular breakpoint	Genetics[a]
In(3L)D^{1}	69D3-4; 70D1	+2 to +5	DV
Ab(3L)D^{3,b}	69D3-4; 70D1 + coding deletion	+2 to +5 and internal	N
T(2; 3)D^{4}	21D1; 70C15-D2	−9.8 to −12.4	D, L, H
Df(3L)D^{5}	70D2; 70D5	+2 to +4	D, V
In(3)D^{6}	70C14-D1; 91A3-8 on	−6.5 to −9	D, L, H
T(2; 3)D^{7}	32DE; 70D1-3 on	−16 to −18	D, L, H
In(3L)D^{r8}	70D1-2; 79E2	−5.5 to −6.5	L, H
In(3L)D^{10}	70A1-2; 70D1-2	−21 to −24	D, L, H
D^{r321}	ND	+1 to +1.5	L, H

[a] The phenotypic class of each allele. D, exhibits the dominant wing phenotype; V, viable in combination with null alleles; L, lethal in combination with null alleles; H, hypomorphic allele designated due to weak segmentation phenotype; N, protein null.

[b] The *D*^{3} allele is complex, is a spontneous partial revertant of *D*^{1}, and retains the *D*^{1} inversion; however, it has suffered an additional DNA lesion within the coding sequence of the *Dichaete* transcription unit. There is no detectable Dichaete protein and the embryonic phenotype is amorphic.

development (Nambu *et al.*, 1990; Russell *et al.*, 1996). In order to examine *Dichaete* function in developmental processes other than segmentation we utilized a series of alleles associated with chromosomal rearrangements in the vicinity of the *Dichaete* transcription unit. Six of the alleles (*D*^{4}, *D*^{6}, *D*^{7}, *D*^{r8}, *D*^{10}, and *D*^{321}) are recessive lethal in combination with *Dichaete* null alleles. Four of these alleles (*D*$^{4,\ -6,\ -7,\ -10}$) are associated with a dominant wing hinge phenotype which results from ectopic expression of *Dichaete* in the wing imaginal disks, suggesting that the mutations retain functional Dichaete protein (Russell *et al.*, 1996; S. Russell, unpublished data). An additional two alleles (*D*^{1} and *D*^{5}) have similar dominant wing phenotypes but are viable in combination with lethal alleles (Table 1). The molecular mapping of the breakpoints associated with these chromosomes shows that all but one of the lethal alleles map 3′ to the *Dichaete* transcription unit whereas the two viable alleles map 5′ (Fig. 1). The *D*^{r321} lethal allele, generated by imprecise excision of a P-element close to the 5′ end of the transcription unit, is associated with a deletion of approximately 200 bp of DNA 1 kb upstream of the *Dichaete* translation start. The 5′ extent of the lethal complementation group is defined by the *D*^{5} breakpoint at +3.5 kb, since this allele is viable with null alleles. The breakpoints of the 3′ alleles, *D*$^{r8,\ -6,\ -4,\ -7,\ -10}$, map downstream of the *Dichaete* polyadenylation site to a region spanning some 25 kb (Fig. 1). Since the most distal of those, *D*^{10}, is a lethal allele further sequences distal to −25 kb are required for *Dichaete* function and we have yet to map the 3′ extent of the gene.

In combination with a *Dichaete* deletion all of the new alleles show segmentation phenotypes; however, the defects are restricted to one or two segments and they are considerably weaker and less frequent than the phenotypes observed with null alleles (Nambu and Nambu, 1996; Russell *et al.*, 1996). Thus, the *Dichaete* product appears to be intact and instead the new alleles uncover other vital *Dichaete* functions. Taken together, the mapping of the various alleles indicates that the *Dichaete* locus spans at least 30 kb of DNA. Since the intronless *Dichaete* transcription unit covers only 1.8 kb of DNA (Nambu and Nambu, 1996; Russell *et al.*, 1996), this suggests that extensive regulatory sequences, mostly located 3′ to the transcription unit, are essential for wild-type *Dichaete* function.

Dichaete Has Extensive 3′ Regulatory Sequences That Control Tissue-Specific Expression

To characterize the phenotypes associated with the *Dichaete* alleles, and to determine whether the sequences removed by the chromosomal aberrations are indeed necessary for correct *Dichaete* expression, we carried out a developmental analysis of *Dichaete* expression with each of the alleles in hemizygous condition. In the wild-type, *Dichaete* expression is dynamic and widespread throughout embryogenesis (Nambu and Nambu, 1996; Russell *et al.*, 1996; Ma *et al.*, 1998; Sánchez-Soriano and Russell, 1998) and we have focussed our analysis on the most prominent domains; the early blastoderm, the developing CNS and the hindgut. In the wild-type embryo at stage 5, Dichaete is found in a broad central domain in the trunk region and an anterior domain in the procephalic region. In all of the 3′ alleles, central domain expression is present, albeit reduced, compared to wild type (Figs. 2A–2D). This is not unexpected since all of these alleles exhibit almost normal segmentation phenotypes. In the 5′ allele, *Dr*^{321}, central domain expression is indistinguishable from wild type. Taken together these data indicate that central domain regulatory elements lie proximal to the *D*^{r8} breakpoint at −6 kb but distal to the *D*^{r321} lesion.

In the developing CNS there are three prominent sites of

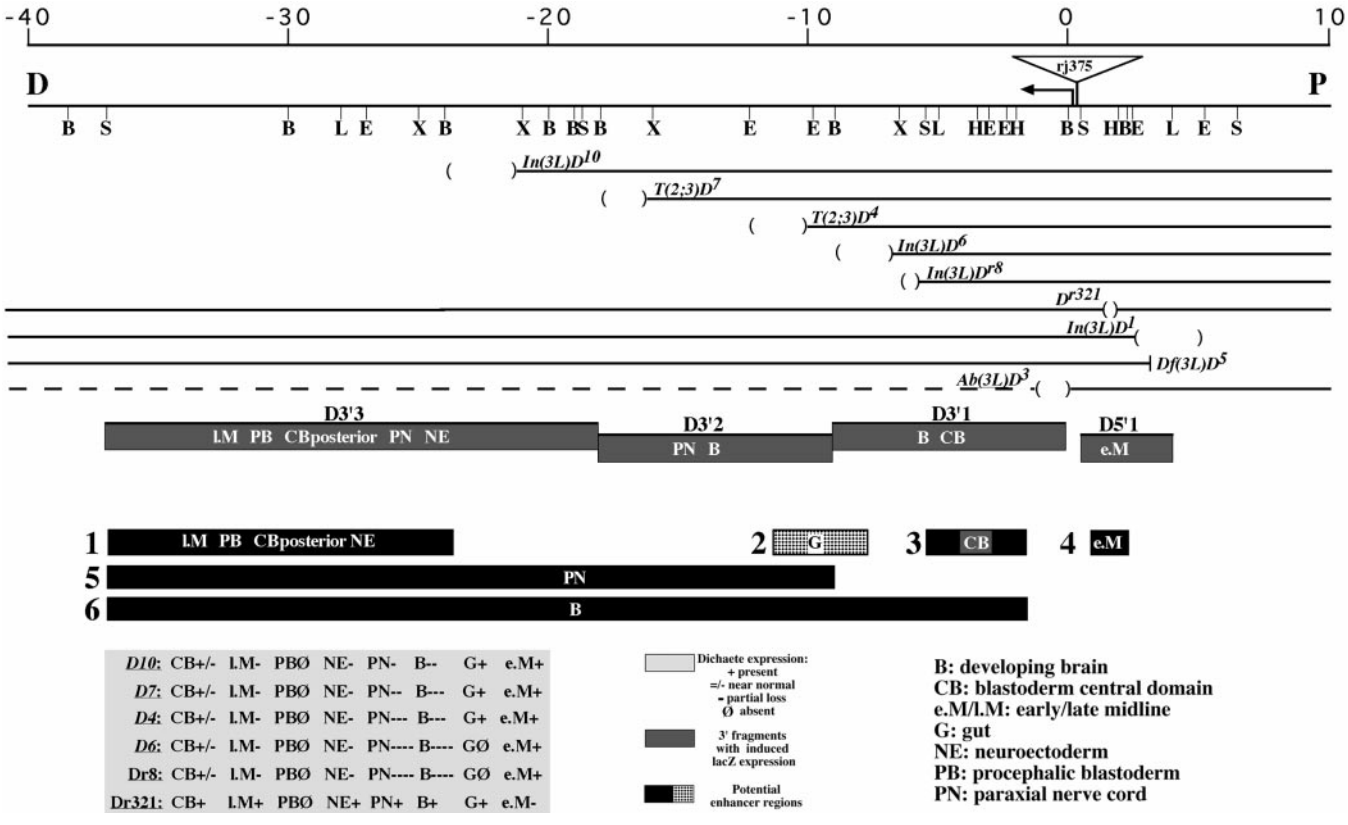


FIG. 1. Molecular analysis of the *Dichaete* region at 70D1. The top line is a scale bar in kilobases, which is zeroed at the start of *Dichaete* translation; the distal (D) and proximal (P) ends with respect to the centromere are indicated. A solid arrow marks the extent of the Dichaete transcription unit and the insertion site of the P-element p[PZ]rj375 (Nambu and Nambu, 1996) is indicated. The next line is a restriction map of the *Dichaete* region indicating restriction sites of the following enzymes: B, *Bam*HI; E, *Eco*RI; H, *Hin*dIII; L, *Sal*I; S, *Sac*I; X, *Xho*I. Parentheses below indicate restriction fragments to which breakpoints of each allele (added in italics) were mapped; adjacent lines indicate the respective remaining DNA. In the case of D^3 the dotted line indicates uncertainty in the nature of the secondary aberration which disrupts the coding sequence. The light gray box lists for the recessive lethal alleles D^{10} to D^{r321} which subpattern of *Dichaete* expression is lost, respectively (for explanation of letter code used in light gray, middle gray, and black boxes see box at the bottom). D^1 and D^5 (a straight line indicates the precise location of the breakpoint of D^5) are viable alleles; D^3 is a protein null. Middle gray boxes represent the extent of the DNA fragments used to construct the lacZ reporter constructs D3′1–3 and D5′1; white letters indicate the subpattern of Dichaete expression as revealed by lacZ-reporter gene expression. Black and stippled boxes indicate the inferred location of enhancer regions for subpatterns of *Dichaete* expression: Box 1 is suggested through loss of expression in D^{10} and lacZ-expression in D3′3, box 2, through loss of expressing in D^6 but not D^4 (this box is stippled because although there is a lack of expression in the mutants no constructs express in the gut); box 3, through the presence of Dichaete expression in D^{r8}, lacZ expression in D3′1, and the distal end of the transcription unit; box 4, through presence of Dichaete or lacZ expression in D^1 and D5′1 and absence of *Dichaete* in D^{r321}; box 5, through decrease of Dichaete expression from D^{10} to D^6 and presence of lacZ expression in D3′2 and D3′3; and box 6, through decrease of Dichaete expression from D^{10} to D^{r8} and induction of lacZ expression by all 3′ Dichaete fragments.

Dichaete expression, the ventral midline, the CNS of the trunk outside the midline (from now on referred to as the "paraxial nerve cord"), and the cephalic CNS. Expression appears to be independently regulated in each of these domains. In the wild type, *Dichaete* is expressed in the midline from stages 7 to 14 (Sánchez-Soriano and Russell, 1998). We identified two groups of alleles that have opposite effects on midline expression: The 5′ mutation, D^{r321}, lacks midline expression until stage 12 and thereafter it appears to be wild type (not shown). This suggests that regulatory elements for early midline expression are located close to the 5′ end of *Dichaete*. Accordingly, all of the 3′ alleles show normal midline expression up to stage 12 (Figs. 2I–2L). However, subsequently midline expression in all of the 3′ alleles declines prematurely, indicating regulatory elements for late midline expression distal to −25 kb.

In the paraxial nerve cord of wild-type embryos, Dichaete is detected in the neuroectoderm from stage 7 and subsequently in a complex and dynamic pattern in many segregating neural precursors and their progeny. All the 3′

breakpoint alleles retain substantial (though less than wild type) early neurectodermal expression until stage 10. By stage 11 this expression is affected; the alleles with breakpoints proximal to −12 kb ($D^{r8,\ -6,\ -4}$) show a strong reduction in paraxial nerve cord expression (Figs. 2K and 2L). In mutant embryos with breakpoints distal to −12 kb ($D^{7,\ -10}$) Dichaete product can be detected in segregated neuroblasts and their progeny; in D^7 only a single strongly staining cell and a few weakly staining cells are found in each hemisegment; and in the more distally mapping D^{10} allele more cells are stained, though still less than in wild type (Fig. 2J). Taken together, these data suggest several discrete regulatory regions for paraxial nerve cord CNS expression: First, an early neurectodermal element proximal to the D^{r8} breakpoint. Second, elements in the area distal to the D^{r8} breakpoint up to beyond the D^{10} breakpoint, which drive expression in different sets of segregated neuroblasts and progeny. Third, sequences distal to the D^{10} breakpoint, which drive late neurectodermal expression together with expression in further neuroblasts and progeny.

At the anterior end of the wild-type embryo *Dichaete* is expressed at the blastoderm stage in a domain in the procephalic region. This expression is normal in D^{r321}, but lacking in all of the mutant alleles with 3′ breakpoints, indicating that regulatory elements for procephalic blastoderm expression lie distal to −25 kb (Figs. 2A–2D). Subsequently, *Dichaete* is expressed, again in a complex and dynamic pattern, in the developing brain. Much of the postblastoderm brain expression is lost in the 3′ alleles (but not in D^{r321}). The exception is a patch of expression in the dorsal region of the brain that increases in intensity in alleles with increasing portions of 3′ DNA (Fig. 2). This suggests that there may be regulatory elements required for brain expression dispersed throughout the 3′ region.

Finally, in wild-type embryos, *Dichaete* is expressed in the developing hindgut from stage 10. In *Dichaete* mutant embryos there is a sharp demarcation between D^4, which has wild-type expression, and the alleles with breakpoints proximal to this (D^6 and D^{r8}) showing no discernible expression (Figs. 2O and 2P). This indicates that a hindgut regulatory element is located proximal to −10 kb.

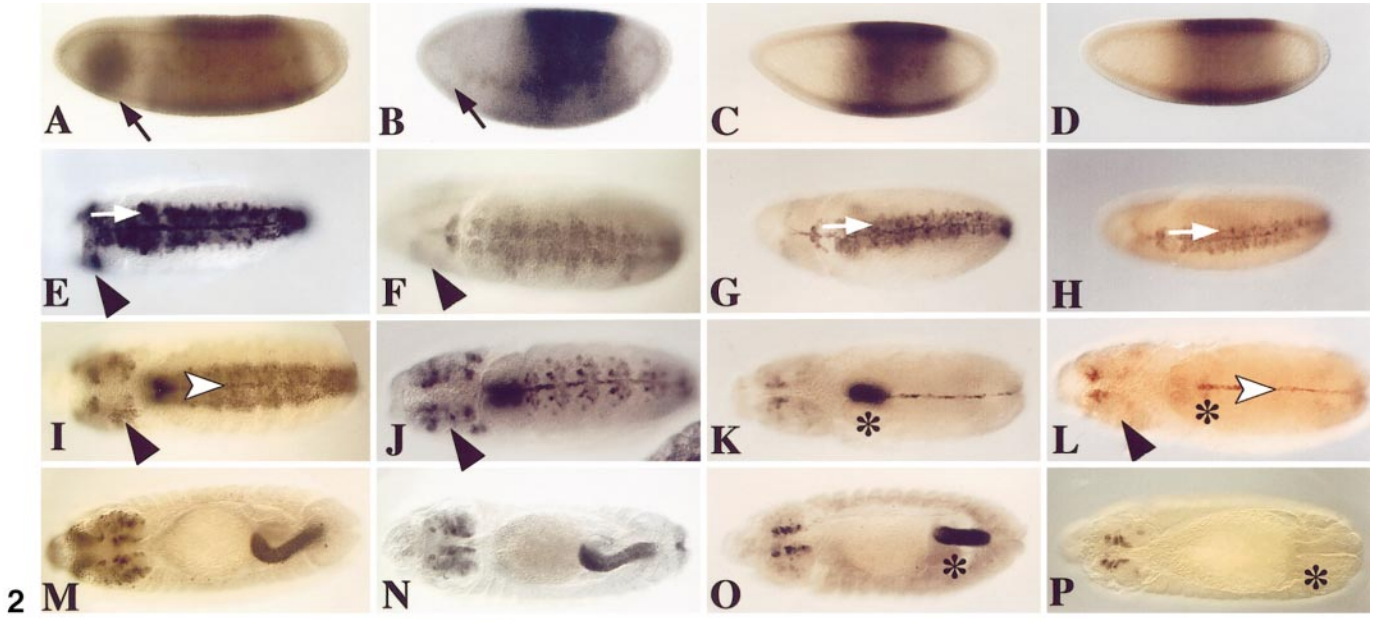
Flanking Sequences Drive Reporter Gene Expression in Dichaete Domains

We tested the suggested organization of the regulatory region of *Dichaete* by generating transgenic flies which express the *lacZ* gene under the control of sequences flanking the *Dichaete* transcription unit. We generated four constructs, one encompassing the 5′ extent of the gene from the D^5 breakpoint to the start of translation (called D5′1) and three which span the 3′ region (called D3′1, D3′2, D3′3; for their location see Fig. 1). In all cases we confirmed that reporter gene expression corresponds with wild-type *Dichaete* expression by double labeling with antibodies against β-galactosidase and Dichaete.

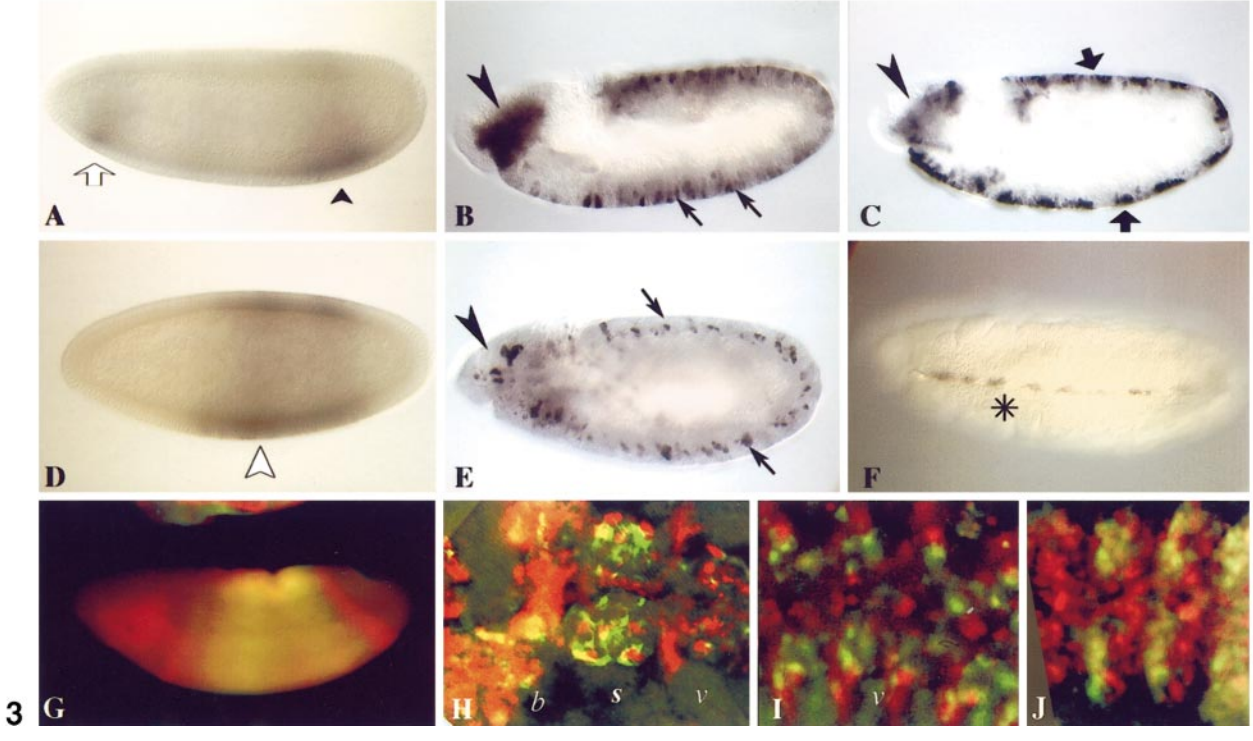
The 5′ construct directs reporter expression in the midline from stage 10 (Fig. 3F). The proximal D3′1 construct shows reporter expression in the central domain of the early blastoderm and from stage 10 in a subset of cells in the developing brain. The central domain reporter gene expression does not extend as far anterior or posterior as wild-type *Dichaete* suggesting additional regulatory elements in other areas (Figs. 3D and 3G). The middle 3′2 construct drives reporter expression in a subset of *Dichaete*-expressing cells in the paraxial nerve cord and the brain (Figs. 3E, 3H, and 3I). Finally, the distal 3′3 construct shows expression in the procephalic region at blastoderm as well as a posterior stripe which corresponds to the blastodermal central domain missing in the proximal 3′ construct (Fig. 3A). In addition, the 3′ construct shows reporter coexpression with a subset of *Dichaete*-expressing cells in the neurectoderm, paraxial nerve cord, and, from stage 12, in the midline (Figs. 3B, 3C, and 3J). The reporter gene expression patterns observed with each of the constructs correlate with the domains of *Di-*

FIG. 2. The expression of Dichaete is altered in the 3′ regulatory alleles. *Dichaete* expression during embryonic development revealed by staining with anti-Dichaete antiserum in different mutant backgrounds. Wild type (A, E, I, and M); $D^{10}/Df(3L)GS1a$ (B, F, J, and N); $D^4/Df(3L)GS1a$ (C, G, K, and O); $D^{r8}/Df(3L)GS1a$ (D, H, L, and P). (A–D) Lateral views of stage 5 embryos showing expression in the trunk region; there is no staining in the procephalic region in *Dichaete* mutant embryos (small black arrow). (E–H) Ventral views of stage 10 embryos; *Dichaete* staining in the developing brain is absent in *Dichaete* mutant embryos (black arrowhead) and the trunk neurectodermal expression is reduced (small white arrow). (I–L) Dorsal views of stage 11 embryos; expression in the midline appears wild type in all of the alleles (white arrowhead). Expression in the developing hindgut is normal in D^4 and D^{10} embryos (J and K) but absent in D^{r8} (L) (black asterisk). In addition, there is no staining in the paraxial trunk CNS in D^4 or D^{r8} (K and L) and in D^{10} expression is less than wild type. In the dorsal brain, D^{10} exhibits a more wild-type pattern than the other alleles (black arrowhead). (M–P) Dorsal views of stage 13 embryos. Hindgut expression is still absent from D^{r8} and brain expression is reduced in all of the alleles compared to wild type.

FIG. 3. Sequences flanking the *Dichaete* transcription unit drive β-galactosidase expression in a subset of *Dichaete* domains. (A–F) Embryos containing *LacZ* reporter constructs stained with anti-β-Gal –antibody; for a comparison with the pattern of *Dichaete* expression see Fig. 1. (G–J) Embryos containing *LacZ* reporter constructs stained with anti-Dichaete (red) and anti-β-Gal (green) antibodies and visualized by confocal microscopy; Dichaete is predominantly localized in nuclei while the β-gal is localized cytoplasmically. (A–C and J) The D3′3 construct drives *LacZ* expression from early stage 5 in an anterior spot in the head region (white arrow in A) and in a posterior domain of the trunk (small arrowhead in A); at stage 10 (B, lateral view) and stage 11 (C, lateral view) expression is detected in the



2



3

neuroectoderm and in neural cells of the cephalic region (large arrowhead), in segregated neural cells of the trunk (thin arrow), and from stage 11, in the neuroectoderm of the trunk (flat arrow in C); in flat preparations of stage 13 embryos (J, ventral view) colocalization with Dichaete is detected in the developing nervous system of the trunk. (D and G) The D3′1 construct expresses *LacZ* from stage 5 in a central domain (white arrowhead in D; yellow staining in G) which does not extend as far anterior or posterior as wild-type Dichaete (note red regions in G). (E, H, and I) The D3′2 construct labels, at stage 11 (E, lateral view), neural cells in the trunk (thin arrows) and in the developing brain (long arrowhead); at stage 13 (H and I, ventral views of dissected embryos) many cells coexpress β-galactosidase and Dichaete in brain (*b*), subesophageal ganglion (*s*), and ventral nerve cord (*v*). (F) The D5′1 construct drives *LacZ* expression in the midline (asterisk; stage 11 embryo in ventral view).

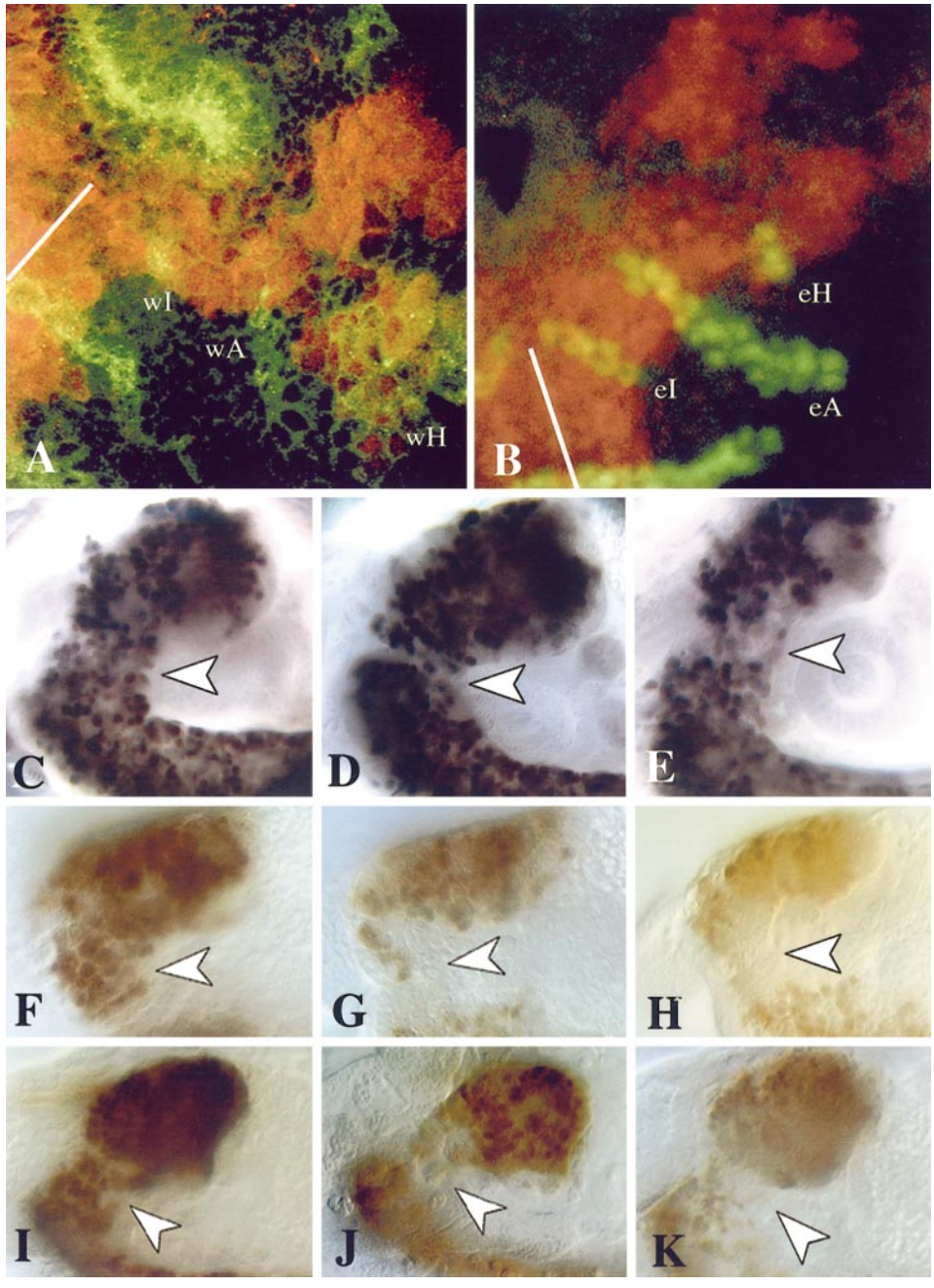


FIG. 4. *Dichaete* is required during brain development. (A and B) Flat preparations of late stage 11 embryonic heads stained with anti-Dichaete (red) and anti-Wingless (green) in A or anti-Dichaete (red) and anti-Engrailed (green) in B visualized by confocal microscopy. Dichaete coexpresses with Wg and En (yellow) in the intercalary spot (wI and eI), a subset of antennal stripe cells (wA and eA), the Wg head blob (wH), and the En head spot (eH). The white lines indicate the position of the midline. (C–E) Zfh-1 expression in stage 16 embryonic brains of wild type (C), $D^{r8}/Df(3L)GS1a$ (D), and $D^{3}/Df(3L)GS1a$ (E); the arrowheads indicate the region of the brain with a gap in expression. (F–K) *castor* expression monitored with the *ming LacZ* line at stage 12 (F–H) and stage 16 (I–K) in wild type (F and I), $D^{r8}/Df(3L)GS1a$ (G and J), and $D^{3}/Df(3L)GS1a$ (H and K). The loss of LacZ-expressing cells in the brain is indicated with the arrowheads.

chaete expression lost in the regulatory alleles (for details see Fig. 1; for gut expression also see discussion). This indicates that the regulatory alleles identify *bona fide* *Dichaete* regulatory sequences and show that *Dichaete* is regulated by enhancer elements dispersed throughout a 30-kb region.

Using Regulatory Alleles to Probe Tissue-Specific Dichaete Function in the Midline and Paraxial Ventral Nerve Cord

Since our analyses have revealed a set of *Dichaete* mutant alleles that affect tissue-specific aspects of *Dichaete* expression, these alleles can be used to study *Dichaete* requirements in different *Dichaete*-expressing tissues. As we have previously described, *Dichaete* is required for the correct specification and differentiation of midline glia, and loss of *Dichaete* in the midline results in characteristic defects in the mature axon scaffold (Sánchez-Soriano and Russell, 1998). In agreement with this, D^{r321} mutant embryos, which lack midline expression before stage 12/13, have defects in the development of thoracic and posterior abdominal midline glia and of the neuropile (fusion of commissures, thinning and collapse of longitudinal connectives toward the midline; data not shown). The 3′ alleles, which lack only late midline expression of Dichaete, show no defects. This finding confirms our previous work on the requirements of *Dichaete* in the midline and demonstrates that *Dichaete* function is required before stage 12/13. We next examined the effects of loss of *Dichaete* expression in the paraxial nerve cord, in the cephalic CNS, and in the hindgut primordium.

Of the regulatory alleles characterized, D^{r8} has the strongest effect on *Dichaete* expression in the paraxial nerve cord. From late stage 11 there is no expression in segregated neuroblasts and their progeny; however, earlier neurectodermal expression is prominent. We examined the nervous system of D^{r8} hemizygous embryos using global neuronal markers, such as BP102, and the specific lineage markers Even skipped and Engrailed. In all cases the nervous system appeared wild type, similar results were obtained with D^{6} and D^{10} (data not shown). In contrast, embryos carrying *Dichaete* null alleles show defects in Eve lineages (an increase in the number of cells in the position of the EL cluster, the RP2, aCC, and pCC neurons; Sánchez-Soriano, 1999) and a loss of En-expressing cells, a phenotype previously reported for *Dichaete* null alleles (Nambu and Nambu, 1996). The discrepancy between the null and regulatory alleles may reflect the fact that the regulatory alleles retain early neurectodermal expression and it is only this early expression which is important for the phenotypes observed. Alternatively, defects in the paraxial nerve cord of null mutant embryos may be caused secondarily by early segmentation defects (Patel *et al.*, 1989). In this case, *Dichaete* has no easily detectable primary function in the CNS or its function is masked by other redundant neural *Sox* genes, as is believed to be the case in vertebrates (see discussion).

Dichaete Function in the Developing Brain

Dichaete is expressed in the procephalic neurectoderm before neuroblast formation and subsequently in the developing brain. Since the expression of Dichaete in the developing brain is widespread, complex, and dynamic, including neurectodermal cells, neuroblasts, and differentiated neuronal cells, it is difficult to precisely map with respect to the existing neuroblast maps. We have mapped some of the Dichaete-positive cells by double labeling with Dichaete, Wingless (Wg), and Engrailed (En) antibodies. At embryonic stage 11, Wg and En are expressed in adjacent domains called wg/en intercalary spot, wg/en antennal stripe, and wg/en head blob or spot. These domains of expression contribute to the tritocerebrun, deuterocerebrun, and protocerebrun, respectively, and, in particular, En expression can be used as a landmark to delimit these three brain neuromeres (Younossi-Hartenstein *et al.*, 1996; Schmidt-Ott and Technau, 1992; Richter *et al.*, 1998). We find that, at stage 11, Dichaete coexpresses with these markers in the intercalary spot, a portion of the antennal stripe and in the "head blobs" (Figs. 4A and 4B). However, it is apparent from this figure that within the three brain neuromeres, very many brain cells express Dichaete at high levels.

All of the 3′ regulatory alleles lack procephalic neurectodermal expression at the blastoderm stage in addition to most staining in the developing brain until late stage 10. After this D^{r8} mutant embryos show very little staining whereas D^{10} mutant embryos have more extensive expression (Fig. 2). To look for defects in the development of *Dichaete* mutant brains we first examined the expression of *Zfh-2*, a transcription factor expressed in many neuronal lineages (Lai *et al.*, 1991). In comparison to the wild type, in D^{r8} stage 16 mutant embryos, we found a reduction in Zfh-2-positive cells in the posterior part of the brain, most likely the tritocerebrun (Figs. 4C–4E). Similarly, *ming-LacZ* expression, another marker abundant in the CNS (Kambadur *et al.*, 1998), is reduced in this region at stage 12 and 16 when analyzed in null and D^{r8} mutant backgrounds (Figs. 4F–4K). However, anti-Fasciclin II stainings of wild-type and mutant embryos reveal no detectable differences in the location or arrangement of the major brain commissures and longitudinal connectives (Sánchez-Soriano, 1999). Thus, despite widespread expression, *Dichaete* requirement is mainly restricted to a specific region of the embryonic brain and has no apparent function in the development of the major identified axon tracts in the brain.

The tritocerebrum originates from neuroblasts that segregate from the intercalary segment during stages 10 and 11 (Younossi-Hartenstein *et al.*, 1996). The intercalary segment is marked by the expression of the homoeotic gene *labial* (*lab*) and, as development proceeds, *labial* expression is maintained in the tritocerebrum (Hirth *et al.*, 1998). In D^{r8} hemizygotes at stage 16 we found a reduction in the number of *lab*-expressing cells compared to the wild type (Figs. 5C and 5D: 46 vs 29, $n = 5$); D^{10} and the null allele D^{3} show similar phenotypes. In the wild-type brain, *Dichaete* is coexpressed in a subset of *lab*-expressing cells until stage 11 but not thereafter (Figs. 5A and 5B). Therefore, *Dichaete* appears to be required early for the correct development of this fraction of *labial*-expressing cells. We next examined earlier stages of brain development. Preliminary analysis of *lethal of scute* (*l'sc*)-expressing cells in stages 10–11 D^{r8}

mutant embryos, revealed widespread defects in the tritocerebrum anlage (not shown). This is in contrast to other regions of the brain which appear unaffected or show subtle defects (e.g., deutocerebrum). Since the *l'sc* pattern of expression is complex and dynamic, we used the more restricted markers Wg and En. At stage 10 in the wild-type embryo, there are one to two large Wg-positive cells (most likely neuroblasts) in the intercalary segment. Wg expression in these cells is undetectable or severely reduced in D^{r8} hemizygous embryos, whereas other Wg domains in the head, which normally express Dichaete (Fig. 4A), appear unaffected (Figs. 5E and 5F). Similar data were obtained in embryos expressing *lacZ* under the control of the *wingless* promoter (not shown). In the wild type, at stage 11, En is expressed in at least 10 neuroblasts and/or neuronal cells in the intercalary segment. In D^{r8} mutant embryos En expression in the intercalary segment is strongly reduced since four or fewer positive cells are detected. As for Wg, the other En domains in the developing brain appear to be unaffected (Figs. 5G and 5H). Taken together these data indicate that *Dichaete* is required for the development of a subpopulation of cells, including neuroblasts, within the anlage of the tritocerebrum at stage 10/11 when the neural precursors of the tritocerebrum form (Schmidt-Ott and Technau, 1992; Younossi-Hartenstein *et al.,* 1996). Therefore, loss of *Dichaete* might affect the correct specification or differentiation of specific neuroblast lineages, leading to loss of *Zfh-2, ming-lacZ,* and *lab* expression.

Dichaete in the Hindgut

Dichaete is expressed from stage 10 in the hindgut primordium and from stage 13 to stage 16 it is restricted to the ectodermal layer of the large intestine; the central region of the hindgut (Figs. 6A and 6M). During normal hindgut development a series of cell divisions, which cease at stage 11, generate the cells of the hindgut epithelium. After germ band shortening, an elongation phase involving cellular reorganization in the large intestine, completes morphogenesis (Skaer, 1993). In D^{r8}, which shows no detectable *Dichaete* expression in the gut at any stage, anti-Engrailed (En) staining suggests that the hindgut is normal until stage 13 since approximately the same number and distribution of cells are observed in wild type and mutant (not shown). Subsequently, the hindgut of D^{r8} mutant embryos is shorter, its lumen wider and En expression is reduced in the large intestine when compared to wild-type embryos (Figs. 6B–6D; similar defects are found in *Dichaete* null mutant embryos). Thus, the appearance of defects in *Dichaete* mutant hindguts correlates with the elongation phase of normal hindgut development, suggesting that *Dichaete* is required for the second step of hindgut growth. It has been proposed that the elongation phase of hindgut morphogenesis is driven by cellular reorganization and cell shape changes. For example, at stage 16, the cells of the wild-type gut have adopted a columnar shape when observed with the cytoskeletal marker anti-Spectrin (Blake *et al.,* 1998; Campos-Ortega and Hartenstein, 1997; Jack and Myette, 1997; Pesacreta *et al.,* 1989). We examined *Dichaete* mutant embryos with anti-Spectrin and found, at stage 16, that the cells of the large intestine are short and rounded in contrast to the thin columnar cells of the wild type (Figs. 6E and 6F). Thus, the defects in gut morphology found in *Dichaete* mutant embryos may be, in part, a consequence of defects in cell shape change.

The formation of the *Drosophila* hindgut requires a cascade of gene activities. The terminal gap genes *tailless* and *caudal* activate expression of genes, including *forkhead* and *brachyenteron,* which act as key regulators in the establishment and maintenance of the hindgut primordium (Diaz *et al.,* 1996; Wu and Lengyel, 1998). In the differentiating hindgut, *wg, hedgehog* (*hh*), and *decapentaplegic* (*dpp*) are expressed in restricted domains in the hindgut epithelium which correspond to morphological structures (small and large intestine, rectum). It has been suggested that the activities of these signaling molecules mediate subdivision and domain-specific development of the hindgut epithelium (Hoch and Pankratz, 1996). *dpp* is expressed throughout the central region of the hindgut in the large intestine, which is flanked by the small intestine anteriorly and the rectum posteriorly. Adjacent and nonoverlapping *hh* and *wg* expression domains mark both small intestine and rectum. The *hh* domains are both adjacent to the *dpp* domain and the *wg* domains mark the anterior and posterior limits of the hindgut. The expression of these three genes was analyzed in null and D^{r8} mutant embryos. We found no effects on either of the *wg* domains or on the anterior *hh* domain; however, the posterior domain of *hh* expression expands into the large intestine and *dpp* expression is repressed (Figs. 6G–6K). Thus the hindgut phenotype observed in *Dichaete* mutants could be due to an expansion of *hh* expression or to the loss of *dpp* expression.

To distinguish between these possibilities we used an *en-GAL4* driver line (expressing in most of the hindgut) to ectopically express *UAS-Dichaete, -hh,* and *-dpp* constructs (Brand and Perrimon, 1993). In wild-type flies expressing *Dichaete* under the control of *en-GAL4, Dichaete* is extended posteriorly (Fig. 6N). As a consequence, *hh* is partially repressed in the rectum (Fig. 6L), but there are no detectable effects on *dpp* expression. Thus, *Dichaete* can repress *hh* but is not sufficient to induce *dpp.* To test whether the expansion of *hh* is the cause of the hindgut phenotypes, we expressed *hh* under the control of *en-GAL4.* We found no effect on *dpp* expression and no significant defects in hindgut morphology. Similar data have been reported with heat shock driven *hh* (Hoch and Pankratz, 1996). Therefore, the *Dichaete* phenotype might be mediated through loss of *dpp* function. To examine this possibility we expressed *dpp* in the hindgut of *Dichaete* null mutant embryos with *en-GAL4.* We observe significant, though variable, rescue of both hindgut morphology and En expression. Hindgut length in the D^3/Df; UAS*dpp* embryos is 179 ± 33 μM (n = 12) compared to 131 ± 12 μM (n = 14) in D^3/Df alone and 180 ± 17.5 μM (n = 10) in wild-type

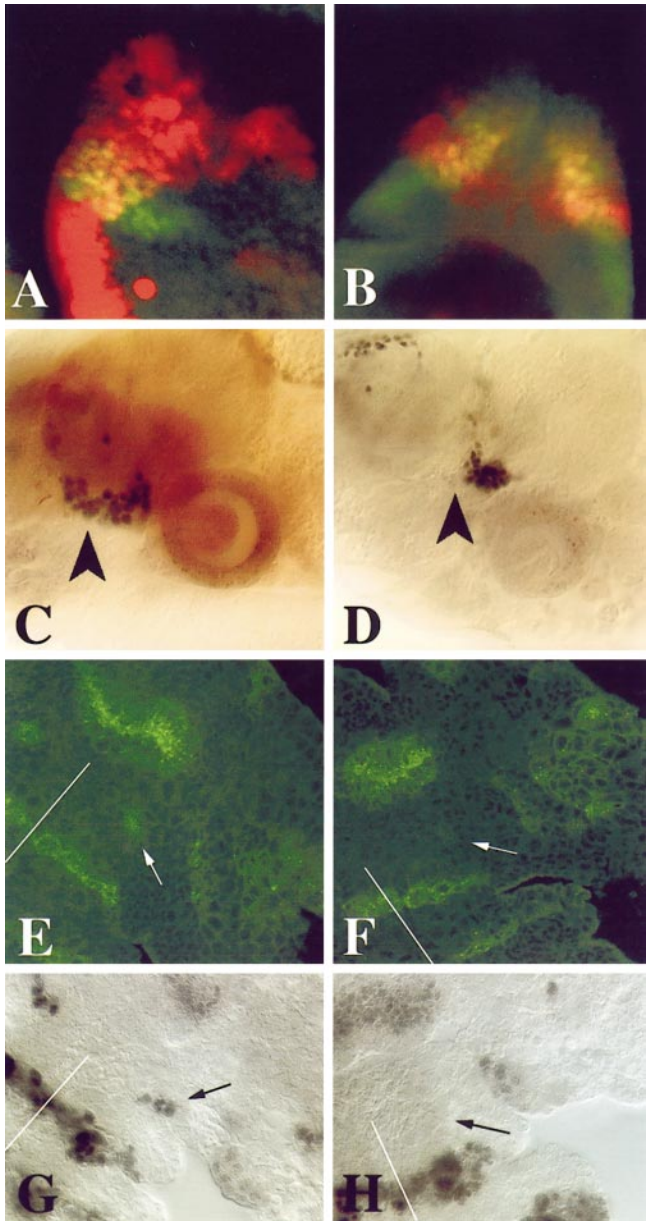


FIG. 5. *Dichaete* mutants have early defects in brain development. (A and B) Expression of Dichaete (red) and Labial (green) in cells of the intercalary segment of stage 10 embryos; E is a lateral view and F is a ventral view. Notice that only a subset of labial cells coexpress Dichaete (yellow). (C and D) Labial-positive cells in the tritocerebrum, marked with a black arrowhead, in lateral views of stage 16 embryos. There is a reduction in the number of Labial-positive cells in $D^{r8}/Df(3L)GS1a$ (D) compared with wild type (C). (E and F) Flat preparation of wild type (E) and $D^{r8}/Df(3L)GS1a$ (F) embryonic heads at stage 10 stained with anti-Wingless; anterior is to the top. White arrows mark segregating neuroblasts within the intercalary segment that have lost wg expression in (F); other wingless domains within the brain are unaffected. The white line marks the position of the ventral midline. (G and H) Flat preparation of wild type (G) and $D^{r8}/Df(3L)GS1a$ (H) embryonic heads at stage 11 stained with anti-Engrailed; anterior is to the top. Black arrows mark segregating neuroblasts and other neuronal cells within the intercalary segment which have lost En expression; other En domains within the brain are unaffected. The white line marks the position of the ventral midline.

(Fig. 6O). Similar rescue is observed when *Dichaete* is expressed in the same way (hindgut length 171 ± 14 μM, $n = 12$; Fig. 6P). Taken together these data indicate that *Dichaete* is required for the correct expression of both *dpp* and *hh* in the hindgut and that the *Dichaete* mutant phenotype is, in part, a consequence of the failure to correctly activate *dpp* expression in the large intestine.

In summary, we have demonstrated here two *in vivo* roles for *Dichaete*, one during embryonic brain development and a second during hindgut morphogenesis. Together with the previously described requirements during segmentation and midline development there is now a set of four different cellular and molecular contexts in which *Dichaete* function can be studied.

DISCUSSION

Dichaete function is required for embryonic gut and brain development, in addition to its functions during segmentation and midline glia cell differentiation previously described (Nambu and Nambu, 1996; Russell *et al.*, 1996; Sánchez-Soriano and Russell, 1998). The mammalian *Sox1, 2* and *3* genes, which are closely related to *Dichaete*, are also widely expressed during embryogenesis, and appear to have multiple developmental functions (Pevny and Lovell-Badge, 1997; Wegner, 1999). In many instances, eukaryotic regulatory molecules are used multiple times during different developmental processes. For example, in *Drosophila*, the *Wnt* gene *wg* is required for patterning the ectoderm and later for gut development, CNS development and imaginal disc patterning (Perrimon and Mahowald, 1987; Wodarz and Nusse, 1998). As a consequence of these multiple functions, an understanding of the direct role of such regulatory molecules in a particular developmental process can be difficult to distinguish from indirect pleiotropic effects (Perrimon, 1998). Thus, in *Dichaete* null mutant embryos, loss of the early segmentation function influences further development of many tissues and complicates the analysis of *Dichaete* function in later embryogenesis. To circumvent this problem we have isolated and characterized a series of regulatory alleles which retain functional Dichaete protein, are almost normal for segmentation functions, result in the tissue-specific loss of *Dichaete* expression, and therefore facilitate the analysis of new *Dichaete* functions.

Sox protein function appears to be context dependent when assayed *in vitro* and thus different *in vivo* requirements may well involve different modes of action (ranging from bending DNA to cofactor-mediated or independent,

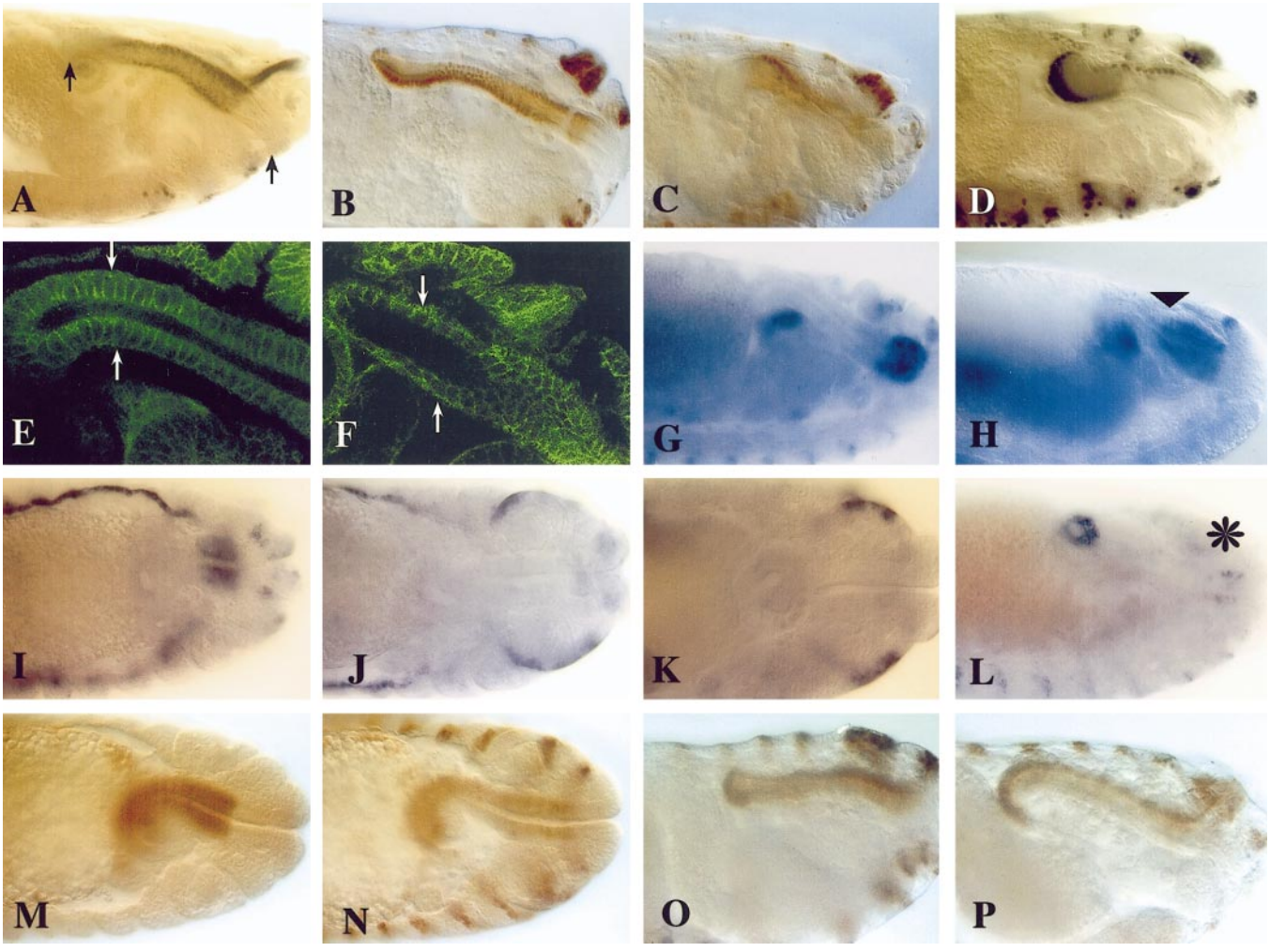


FIG. 6. *Dichaete* mutants have defects in hindgut development. (A–D) Lateral view of the posterior half of stage 16 embryos. (A) Dichaete is restricted to the large intestine as revealed with anti-Dichaete antisera; small black arrows mark the extent of the hindgut. (B–D) Anti-Engrailed staining reveals the hindgut in wild type (B), $D^{r8}/Df(3L)GS1a$ (C), and $D^{3}/Df(3L)GS1a$ (D). Note that the hindgut of *Dichaete* mutants is shorter, the lumen is wider, and Engrailed staining is reduced. (E and F) Lateral views of anti-Spectrin staining in wild type (E) and $D^{3}/Df(3L)GS1a$ (F) stage 16 embryos visualized by confocal microscopy; the *Dichaete* mutant cells are rounded and disorganized compared to the wild type (white arrows). (G and H) Lateral view of *in situ* hybridizations with a *hh* cDNA probe to wild-type (G) and $D^{r8}/Df(3L)GS1a$ (H) stage 13 embryos; in *Dichaete* mutants the posterior *hh* domain expands anteriorly (black arrowhead). (I–K) Dorsal views of *in situ* hybridizations with a *dpp* probe to wild-type (I), $D^{r8}/Df(3L)GS1a$ (J), and $D^{3}/Df(3L)GS1a$ (K) stage 13 embryos; *dpp* expression is absent in the hindgut of *Dichaete* mutants. (L) Lateral view of a stage 13 *en*GAL4/UAS*Dichaete* embryo hybridized with a *hh* cDNA probe; the posterior domain of *hh* is repressed (asterisk). (M and N) Dorsal view of stage 13 embryos stained with anti-Dichaete in wild type (M) and *en*GAL4/UAS*Dichaete* (N); note the posterior expansion of Dichaete. (O and P) Lateral view of a stage 16 embryos stained with anti-Engrailed in *en*GAL4/UAS*dpp;*$D^{3}/Df(3L)GS1a$ (O) and *en*GAL4/UAS*Dichaete;*$D^{3}/Df(3L)GS1a$ (P). In both cases the morphology of the hindgut is partially rescued (the hindgut is thinner and narrower) and engrailed expression is stronger compared with (C).

positive or negative control of transcription; Pevny and Lovell-Badge, 1997; Wegner, 1999). A good example of this is the different activities of Sox2 on the expression of the genes described in the introduction. In the case of *Dichaete*, potential target genes or interacting partners appear to be distinct in different *in vivo* contexts. For example, *Dichaete* regulates the expression of *slit* in the midline (Sánchez-Soriano and Russell, 1998), pair-rule genes during segmentation of the blastoderm (Nambu and Nambu, 1996; Russell *et al.*, 1996) and *dpp* and *hh* during hindgut development (see below). In the midline, Dichaete appears to functionally interact with the POU-domain protein Ventral Veinless (Sánchez-Soriano and Russell, 1998). Similarly, during segmentation it may interact with the POU proteins encoded

by the *pdm-1, −2* genes (Ma *et al.,* 1998). The results presented here identify further potential targets for comparative analysis of Sox function and below we discuss our present insights into *Dichaete* function during gut and nervous system development.

Dichaete in the Hindgut

To our knowledge *Dichaete* is the first transcription factor that is described to be regionally restricted in the hindgut and therefore it may play a role in mediating hindgut regionalization. So far *wg, hh* and *dpp* have been proposed to be involved in the generation of regional specificity. In this model *wg* and *hh* are required in the small intestine and rectum and *dpp* is required in the large intestine. All three of these genes require the activity of *forkhead,* which is expressed throughout the length of hindgut, however, the way in which *wg, hh* and *dpp* are restricted to subdomains is unclear since they do not appear to establish cross regulatory interactions (Hoch and Pankratz, 1996). We find that in the absence of *Dichaete, hh* expression expands into the large intestine and when *Dichaete* is ectopically expressed in the rectum *hh* is partially repressed. This suggests that, directly or indirectly, *Dichaete* is required for the restriction of *hh* to the small intestine. We also find that *dpp* expression is lost from the hindgut of *Dichaete* mutant embryos. The loss of *dpp* expression is apparently not due to the expansion of *hh,* because ectopic expression of *hh* in the large intestine, driven by heat shock (Hoch and Pankratz, 1996), or GAL4 (our own results) does not repress *dpp.* Thus *Dichaete* appears to be required independently for the appropriate expression of both *hh* and *dpp.*

In support of the view that Dichaete regulates *hh* and *dpp* directly, we have examined the available sequence from the regulatory regions of both of these genes for potential Dichaete binding sites (data not shown; Ma *et al.,* 1998). In the case of *hh* we find multiple Dichaete binding sites upstream of the start site and within the first large intron. Interestingly, 3 of the Dichaete binding sites overlap with potential Fkh binding sites. Since Fkh activates *hh,* Dichaete might repress *hh* by blocking the binding of Fkh. In the case of *dpp* we find an number of potential Dichaete binding sites in the P2 and P3 promoter region and further upstream in the *short vein* region (St. Johnston *et al.,* 1990). Although the details of *dpp* regulation in the hindgut are poorly understood, the *shv* region, in particular the region around P2 and P3 promoters, has been implicated in controlling *dpp* expression in this tissue (Masucci and Hoffmann, 1993). Taken together these observations support, but do not prove, that Dichaete is directly involved in the regulation of *hh* and *dpp.*

The aberrant hindgut observed in *Dichaete* mutant embryos could be a consequence of defects in cell shape change. We find no evidence to suggest that proliferation of the hindgut anlage is affected in *Dichaete* mutants although we cannot eliminate the possibility that some cells are lost due to cell death after stage 13. Interestingly, in *Sox1* mutant mice, eye defects are associated with a failure in the elongation of lens fiber cells (Nishiguchi *et al.,* 1998). In the case of the *Dichaete* mutant hindgut, cell shape defects may be due to loss of *dpp,* since *en*-GAL4 driven *dpp* partially rescues the aberrant morphology of *Dichaete* null mutant hindguts. In addition, *dpp* function has already been shown to be involved in cell shape changes during dorsal closure (Riesgo-Escovar and Hafen, 1997). However, the reported hindgut phenotype of *dpp* null mutant embryos is more severe than those we observe in *Dichaete* mutants (Hoch and Pankratz, 1996). It is possible that there is sufficient residual *dpp* expression in *Dichaete* mutants to allow partial gut morphogenesis. This is not entirely unexpected since none of the potential *Dichaete* target genes identified to date (*eve, hairy, runt* and *slit*) are completely repressed by lack of *Dichaete* (Ma *et al.,* 1998; Russell *et al.,* 1996; Sánchez-Soriano and Russell, 1998). It is also possible that the hindgut defects described in *dpp* null mutant embryos are partly due to pleiotropic effects from *dpp* function during dorsoventral patterning.

Dichaete in the Nervous System

Like its vertebrate homologue *Sox2,* which is expressed from the very early stages of nervous system development in fish, frogs, chicken and mouse (Collignon *et al.,* 1996; Mizuseki *et al.,* 1998; Uwanogho *et al.,* 1995; Virz *et al.,* 1996), *Dichaete* is widely expressed in the developing nervous system of *Drosophila*. Although the temporal aspects of *Dichaete* expression in the developing brain, paraxial ventral nerve cord and neural midline appear analogous, they are regulated independently. The independent regulation of *Dichaete* in different parts of the developing CNS may reflect underlying developmental differences in these tissues. For example, expression of *lethal of scute,* a proneural gene which is required for the competence of neurectodermal cells to become neural precursors, is regulated in different spatial and temporal modes in these three parts of the developing CNS (Younossi-Hartenstein *et al.,* 1996). The regulatory mutant alleles described here allow us to examine the consequences of specific loss of *Dichaete* function in the paraxial ventral nerve cord from stage 11 onwards, in the neural midline and early in the brain. Absence of midline glia cells in D^{r321} mutant embryos, which lack Dichaete expression in the midline up to stage 12, confirms our previous findings that *Dichaete* is required during midline glia development (Sánchez-Soriano and Russell, 1998).

In the absence of *Dichaete* function in the developing brain Zfh-2 and Lab are partially lost in the developing tritocerebrum. Since defects in Wg- and En-positive neuroblasts and neuronal cells of the tritocerebrum are found as early as stage 10, *Dichaete* seems to be required during the process of specification or maintenance of specific neuroblasts. Since *Dichaete* is only expressed until stage 11 in the intercalary region it is probable that the reduced numbers of

tritocerebral cells observed in Dichaete mutants at stage 16 reflects impairment of neuroblast subpatterns and secondary loss or misspecification of daughter cells. In the paraxial nerve cord *Dichaete* null mutant embryos exhibit restricted defects in Eve and En expressing neurons (Nambu and Nambu, 1996; Sánchez-Soriano, 1999), but so far our attempts (e.g. through timed expression of heat shock driven *Dichaete*) have failed to determine unambiguously whether these defects are due to prior requirements for *Dichaete* in segmentation, or to a direct requirement for *Dichaete* in the developing ventral nerve cord. If *Dichaete* plays a direct role in the development of cells in the paraxial ventral nerve cord, it would be likely to execute this function before 11. This is suggested by D^{r8} mutant embryos, which show no obvious defects although they lack virtually all Dichaete expression in the paraxial ventral nerve cord from stage 11 onwards.

The phenotypes we find in the developing brain and paraxial nerve cord are surprisingly subtle given the strong expression of *Dichaete* in these tissues. However, restricted defects have been described in *wg* mutants where defects are only observed in a subset of *wg* expressing brain domains (Richter *et al.,* 1998). Similar observations have also been described for other Sox proteins. For example, mice lacking *Sox1,* which is strongly expressed in the developing nervous system, have no obvious morphological CNS defects (Nishiguchi *et al.,* 1998). In the case of the restricted *Dichaete* brain phenotypes, one explanation is that *Dichaete* might cooperate with a locally restricted cofactor. Since Dichaete expression overlaps with Lab expression in the intercalary segment, both might cooperate in order to maintain the group of cells that are lost in the tritocerebrum of *Dichaete* mutants. Interactions with other transcription factors have been observed for several other HMG-domain proteins. For example, the HMG box of HMG1 interacts with the homeodomain of HOX proteins and Sox2 interacts with the POU protein OCT3/4 (Yuan *et al.,* 1995; Zappavigna *et al.,* 1996). Alternatively, the lack of strong or widespread defects in the brain and paraxial nerve cord of *Dichaete* mutant embryos may be due to redundancy, as has been proposed for vertebrate *Sox1, 2* and *3* (Collignon *et al.,* 1996; Pevny *et al.,* 1998). In support of this view we have recently identified a novel *Drosophila* class B *Sox* gene which is widely expressed in the developing CNS (S.R., unpublished data).

Dichaete Has Complex Regulatory Sequences

Our analyses indicate that the dynamic embryonic expression of *Dichaete* is regulated by a battery of discrete regulatory elements, as indicated by loss of expression domains in regulatory alleles and reporter gene expression driven by *Dichaete* genomic DNA fragments (summarized in Figure 1). The purpose of our analyses was a better understanding and classification of our mutant alleles and, to this end, a better resolution of our rather rough map was not required. One surprising finding is the extent and location of the regulatory elements, the majority of which are located throughout a 25 kb region 3′ to the gene. Extended 3′ regulatory sequences are not unique to *Dichaete,* however, they are apparently not common. One example is the complex expression of *dpp,* which is modulated in part by a series of enhancers located downstream of the polyadenylation signal (St. Johnston *et al.,* 1990). Since we assayed putative *Dichaete* regulatory sequences upstream of a reporter gene it appears that the majority of regulatory sequences act as classical context- and position-independent enhancer elements. However, we cannot eliminate the possibility that we have failed to identify some regulatory elements that need to be located 3′ in order to function. For example, the breakpoint mapping of the hindgut regulatory element shows that it lies between the D^{7} and D^{4} breakpoints, however, neither of the flanking constructs direct reporter expression in the hindgut. This may simply reflect the fact that our constructs are not overlapping and split the hindgut element(s) abolishing its function. It is also possible that the hindgut regulatory sequences are an example of a context-dependant element that needs to be located 3′ for correct function.

ACKNOWLEDGMENTS

We are particularly indebted to S. Desbordes for performing the *ming-LacZ* stainings. We thank the following for providing reagents: J. de Celis, M. Bienz, S. Cohen, A. Hidalgo, and Z. C. Lai. We are grateful to A. Prokop and J. de Celis for critically reading the manuscript and R. Urbach for helpful discussion. This work was supported by an MRC grant to S.R., M. Ashburner, and D. Gubb and BBSRC and EU grants to N.S.S.

REFERENCES

Ambrosetti, D. C., Basilico, C., and Dailey, L. (1997). Synergistic activation of the fibroblast growth factor 4 enhancer by Sox2 and Oct-3 depends on protein–protein interactions facilitated by specific spatial arrangements of factor binding sites. *Mol. Cell. Biol.* **17,** 6321–6329.

Blake, K. J., Myette, G., and Jack, J. (1998). The products of *ribbon* and *raw* are necessary for proper cell shape and cellular localization of nonmuscle myosin in *Drosophila. Dev. Biol.* **203,** 177–188.

Botquin, V., Hess, H., Fuhrmann, G., Anastassiadis, C., Gross, M. K., Vriend, G., and Schöler, H. R. (1998). New POU dimer configuration mediates antagonistic control of an osteopontin preimplantation enhancer by OCT-4 and Sox-2. *Genes Dev.* **12,** 2073–2090.

Brand, A., and Perrimon, N. (1993). Targeted gene expression as a means of altering cell fates and generating dominant phenotypes. *Development* **118,** 401–415.

Campos-Ortega, J. A., and Hartenstein, V. (1997). "The Embryonic Development of *Drosophila melanogaster.*" Springer Verlag, Berlin.

Collignon, J., Sockanathan, S., Hacker, A., Cohen-Tannoudji, M., Norris, D., Rastan, S., Stevanovic, M., Goodfellow, P. N., and Lovell-Badge, R. (1996). A comparison of the properties of *Sox-3*

with *Sry* and two related genes *Sox-1* and *Sox-2*. *Development* **122,** 509–520.

Diaz, R. J., Harbecke, R., Singer, J. B., Pignoni, F., Janning, W., and Lengyel, J. A. (1996). Graded effect of *tailless* on gut development: molecular basis of an allelic series of a nuclear receptor gene. *Mech. Dev.* **54,** 119–130.

Diaz-Benjumea, F. J., and Cohen, S. M. (1994). Wingless acts through the Shaggy zeste-white 3 kinase to direct dorsal–ventral axis formation in the *Drosophila* leg. *Development* **120,** 1661–1670.

Diederich, R. J., Merrill, V. K. L., Pultz, M. A., and Kaufman, T. C. (1989). Isolation, structure, and expression of *labial* a homeotic gene of the Antennapedia Complex involved in *Drosophila* head development. *Genes Dev.* **3,** 399–414.

FlyBase (1998). FlyBase: A *Drosophila* database. *Nucleic Acids Res.* **26,** 85–88.

Foster, J. W., Dominguez-Steglich, M. A., Guioli, S., Kwok, C., Weller, P. A., Stevanovic, M., Weissenbach, J., Mansour, S., Young, I. D., Goodfellow, P. N., Brook, J. D., and Schafer, A. J. (1994). Campomelic dysplasia and autosomal sex reversal caused by mutations in an sry-related gene. *Nature* **372,** 525–530.

Gubbay, J., Collignon, J., Koopman, P., Capel, B., Economou, A., Münsterberg, A., Vivian, N., Goodfellow, P., and Lovell-Badge, R. (1990). A gene mapping to the sex-determining region of the mouse Y chromosome is a member of a novel family of embryologically expressed genes. *Nature* **346,** 245–249.

Hirth, F., Hartmann, B., and Reichert, H. (1998). Homeotic gene action in embryonic brain development of *Drosophila*. *Development* **125,** 1579–1589.

Hoch, M., and Pankratz, M. J. (1996). Control of gut development by *fork head* and cell signaling molecules in *Drosophila*. *Mech. Dev.* **58,** 3–14.

Jack, J., and Myette, G. (1997). The genes *ribbon* and *raw* are required for proper shape of tubular epithelial tissues in *Drosophila*. *Genetics* **147,** 243–253.

Kamachi, Y., Sockanathan, S., Liu, Q. R., Breitman, M., Lovell-Badge, R., and Kondoh, H. (1995). Involvement of sox proteins in lens-specific activation of crystallin genes. *EMBO J.* **14,** 3510–3519.

Kamachi, Y., Uchikawa, M., Collignon, J., Lovell-Badge, R., and Kondoh, H. (1998). Involvement of *Sox1, 2* and *3* in the early and subsequent molecular events of lens induction. *Development* **125,** 2521–2532.

Kambadur, R., Koizumi, K., Stivers, C., Nagel, J., Poole, S. J., and Odenwald, W. F. (1998). Regulation of POU genes by *castor* and *hunchback* establishes layered compartments in the *Drosophila* CNS. *Genes Dev.* **12,** 246–260.

Karess, R. E. (1985). P element mediated germ line transformation of *Drosophila*. *In* "DNA Cloning" (D. M. Glover, Ed.), Vol. II, pp. 121–142. IRL Press, Oxford.

Lai, Z., Fortini, M. E., and Rubin, G. M. (1991). The embryonic expression patterns of *zfh-1* and *zfh-2*, two *Drosophila* genes encoding novel zinc-finger homeodomain proteins. *Mech. Dev.* **34,** 123–134.

Ma, Y., Niemitz, E. L., Nambu, P. A., Shan, X., Goto, T., and Nambu, J. (1998). Gene regulatory functions of Drosophila Fish-hook, a high mobility group domain Sox protein. *Mech. Dev.* **73,** 169–182.

Masucci, J. D., and Hoffmann, F. M. (1993). Identification of two regions from the *Drosophila decapentaplegic* gene required for embryonic midgut development and larval viability. *Dev. Biol.* **159,** 276–287.

Mizuseki, K., Kishi, M., Matsui, M., Shigetada, N., and Sasai, Y. (1998). *Xenopus* Zic-related-1 and Sox-2, two factors induced by chordin, have distinct activities in the initiation of neural induction. *Development* **125,** 579–587.

Nambu, J. R., Franks, R. G., Hu, S., and Crews, S. T. (1990). The single minded gene of *Drosophila* is required for the expression of genes important for the development of CNS midline cells. *Cell* **63,** 63–75.

Nambu, P., and Nambu, J. (1996). The *Drosophila fishhook* gene encodes a HMG domain protein essential for segmentation and CNS development. *Development* **122,** 3467–3475.

Nishiguchi, S., Wood, H., Kondoh, H., Lovell-Badge, R., and Episkopou, V. (1998). SOX1 directly regulates the g-crystallin genes and is essential for lens development in mice. *Genes Dev.* **12,** 776–781.

Patel, N. H. (1994). Imaginal neuronal subsets and other cell types in whole mount *Drosophila* embryos and and larvae using antibody probes. *In* "*Drosophila melanogaster:* Practical Uses in Cell Biology" (L. Goldstein and. E. Fyrbeg, Eds.), Vol. 44, pp. 445–487. Academic Press, New York.

Patel, N. H., Schafer, B., Goodman, C. S., and Holmgren, R. (1989). The role of segment polarity genes during *Drosophila* neurogenesis. *Genes Dev.* **3,** 890–904.

Perrimon, N. (1998). New advances in *Drosophila* provide opportunities to study gene function. *Proc. Natl. Acad. Sci. USA* **95,** 9716–9717.

Perrimon, N., and Mahowald, A. P. (1987). Multiple functions of segment polarity genes in *Drosophila*. *Dev. Biol.* **119,** 587–600.

Pesacreta, T. C., Byers, T. J., Dubreuil, R., Kiehart, D. P., and Branton, D. (1989). *Drosophila* spectrin: The membrane skeleton during embryogenesis. *J. Cell Biol.* **108,** 1697–1709.

Pevny, L. H., and Lovell-Badge, R. (1997). *Sox* genes find their feet. *Curr. Opin. Genet. Dev.* **7,** 338–344.

Pevny, L. H., Sockanathan, S., Placzek, M., and Lovell-Badge, R. (1998). A role for SOX1 in neural determination. *Development* **125,** 1967–1978.

Richter, S., Hartmann, B., and Reichert, H. (1998). The *wingless* gene is required for embryonic brain development in *Drosophila*. *Dev. Genes Evol.* **208,** 37–45.

Riesgo-Escovar, J. R., and Hafen, E. (1997). *Drosophila* Jun kinase regulates expression of decapentaplegic via the ETS-domain protein Aop and the AP-1 transcription factor DJun during dorsal closure. *Genes Dev.* **11,** 1717–1727.

Russell, S., Sánchez-Soriano, N., Carpenter, A. T. C., Wright, C. R., and Ashburner, M. (1996). A SOX-domain encoding gene of *Drosophila melanogaster* functions in embryonic segmentation. *Development* **122,** 3669–3672.

St. Johnston, R. D., Hoffmann, F. M., Blackman, R. K., Segal, D., Grimaila, R., Padgett, R. W., Irick, H. A., and Gelbart, W. M. (1990). Molecular organization of the *decapentaplegic* gene in *Drosophila melanogaster*. *Genes Dev.* **4,** 1114–1127.

Sambrook, J., Fritsch, E. F., and Maniatis, T. (1989). "Molecular Cloning: A Laboratory Manual." Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Sánchez-Soriano, N. (1999). "Analysis of the in Vivo Function of the *Drosophila* Sox Gene *Dichaete*." Thesis, Univ. Cambridge, UK.

Sánchez-Soriano, N., and Russell, S. (1998). The *Drosophila* SOX-Domain protein Dichaete is required for the development of the central nervous system midline. *Development* **125,** 3989–3996.

Schmidt-Ott, U., and Technau, G. M. (1992). Expression of *en* and *wg* in the embryonic head and brain of *Drosophila* indicates a

refolded band of seven segment remnants. *Development* **116,** 111–125.

Sinclair, A. H., Berta, P., Palmer, M. S., Hawkins, J. R., Griffiths, B. L., Smith, M. J., Foster, J. W., Frishauf, A.-M., Lovell-Badge, R., and Goodfellow, P. N. (1990). A gene from the human sex-determining region encodes a protein with homology to a conserved DNA-binding motif. *Nature* **346,** 240–244.

Skaer, H. (1993). The alimentary canal. *In* "The Development of *Drosophila melanogaster*" (M. Bate and A. Martinez-Arias, Eds.), pp. 941–1012. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Tautz, D., and Pfeifle, C. (1989). A non-radioactive *in situ* hybridisation method for the localisation of specific RNAs in *Drosophila* embryos reveals translational control of the segmentation gene *hunchback*. *Chromosoma* **98,** 81–85.

Thummel, C. S., Boulet, A. M., and Lipshitz, H. D. (1988). Vectors for *Drosophila* P-element-mediated transformation and tissue culture transfection. *Gene* **74,** 445–456.

Uwanogho, D., Rex, M., Cartwright, E. J., Pearl, G., Healy, C., Scotting, P. J., and Sharpe, P. T. (1995). Embryonic expression of the chicken sox2, sox3 and sox11 genes suggests an interactive role in neuronal development. *Mech. Dev.* **49,** 23–36.

Virz, S., Joly, C., Boulekbache, H., and Condamine, H. (1996). Zygotic expression of the *zebrafish Sox-19*, an HMG box-containing gene, suggests an involvement in central nervous system development. *Brain Res. Mol. Brain Res.* **40,** 221–228.

Wegner, M. (1999). From head to toes: The multiple facets of Sox proteins. *Nucleic Acids Res.* **27,** 1409–1420.

Wodarz, A., and Nusse, R. (1998). Mechanisms of WNT signaling in development. *Annu. Rev. Cell. Dev. Biol.* **14,** 59–88.

Wu, L. H., and Lengyel, J. A. (1998). Role of caudal in hindgut specification and gastrulation suggests homology between *Drosophila* amnioproctodeal invagination and vertebrate blastopore. *Development* **125,** 2433–2442.

Younossi-Hartenstein, A., Nassif, C., Green, P., and Hartenstein, V. (1996). Early neurogenesis of the *Drosophila* brain. *J. Comp. Neurol.* **370,** 313–329.

Yuan, H., Corbi, N., Basilico, C., and Dailey, L. (1995). Developmental-specific activity of the FGF-4 enhancer requires the synergistic action of Sox2 amd Oct-3. *Genes Dev.* **9,** 2635–2645.

Zappavigna, V., Falciola, L., Citterich, M. H., Mavilio, F., and Bianchi, M. E. (1996). HMG1 interacts with HOX proteins and enhances their DNA-binding and transcriptional activation. *EMBO J.* **15,** 4981–4991.

Received for publication October 12, 1999
Revised January 27, 2000
Accepted January 27, 2000